

Two-tier Match and Load Process of Student Records

The matching weights outlined in this document are putting a higher value on fields that have a greater probability of discriminating matching records. Gender has two characteristics, it is a high accuracy key entry field (only has two possible values) and if a mismatch occurs, this is a high degree of a non-matching record. Subsequently DOB is probably less accurate from a key entry standpoint but by its nature has a higher discrimination value.

The file match process occurs after an incoming file has been validated and is ready for processing against the student master database tables. The purpose of the match process is to ascertain that the student records on the incoming file 1) already exist in the student master database, 2) is a new student record that needs to be added to the student master database, or 3) is an uncertain or ambiguous match and therefore requires human intervention (see match resolution).

In cases 1 and 2 defined above, the student master database is updated immediately. In case 3, the record is "held" in an ambiguous student area that is used by the match resolution process.

Tier 1 Match

The first tier of the match process is to identify the list of potential matching records (currently referred to as the pre-select process). This process will return all records in the student master database that match the incoming record on at least one of the following criteria:

1. Last Name and DOB and Gender
2. Last Name, First Name and Gender
3. First Name, MI, DOB, and Gender
4. STI

Tier 2 Match

The second tier of the match process will process the list of records returned by the tier 1 process and apply the following business rules.

The incoming record contains a unique student identifier

Scan the list of student records returned by the tier 1 process for a matching unique student identifier (note: either zero or one record will be found in the list)

If a matching record is found,

Compare incoming record demographic data to student master demographic data for one of the following conditions:

1. Last Name and DOB match
2. Last Name, First Name and Gender match
3. Local ID, Gender and DOB match

If any of these conditions are met, this is considered a successful full match

Update Record on Student Master

Insert Change History

Increment update student count of File Log

Otherwise this is a potential match situation where the incoming STI matches to an existing STI but the demographic information is different enough to raise a match resolution situation

Set error condition to "match error"

Increment resolution count on File Log

Insert incoming record into the Ambiguous Match table

Insert potential match record into the Potential Match table linking the incoming record with the Student Master record with the same Unique Student Identifier

In addition, continue with the remaining demographic matches as discussed below to determine if there are potential match records that should be considered by

match resolution. This will address the situation where a valid STI was assigned to another student incorrectly in the local systems. This is most often observed when there are duplicate local IDs.

Otherwise a matching record is not found (in this case the incoming record contains an STI that is not currently in the student master database)

This situation indicates that the school system is attempting to send an STI that is in error. This could happen if the user manually updated their local system and miss-keyed the STI (among other potential possibilities)

Set error condition to "match error" to force this record into match resolution

Continue with the remaining demographic matches as discussed below to determine if there are potential match records that should be considered by match resolution

Record Does NOT Contain a Unique Student Identifier or the incoming STI did not result in a matching STI

Weight the student master record returned by the tier 1 process against the incoming record by comparing the following demographic fields

Gender	50
DOB	20
*Last Name	10
**First Name	9
MI	6
Local ID	5
Total	100

When the incoming last name or first name fields compared to the last name, first name in the student master differ based on missing letter(s) or truncation of the names, the following match logic will be applied:

Data: Last Name, First Name, MI, DOB, Gender

Example1: Incoming "SULLENBERGER", "JOHN", "D", "10/12/1980", "M"
Potential "SULLENBERGE" "JOHN", "D", "10/12/1980", "M"
Compare: "SULLENBERGE" to "SULLENBERGE"
Result: Match on Last Name

Example2: Incoming "SULLENBERGER", "JOHN", "D", "10/12/1980", "M"
Potential "SULLENBERGE" "FRED", "D", "10/12/1980", "M"
Do not compare as first names are also different
Result: No Match on Last Name

Example3: Incoming "SULLENBERGER", "JOHN", "D", "10/12/1980", "M"
Potential "SULLENB" "JOHN", "D", "10/12/1980", "M"
Do not compare as the different in lengths is 5
Result: No Match on Last Name

* How last name should be compared for setting weight.

If incoming last name = existing last name, then add weight for last name

Else

If the incoming first name = existing first name

And the incoming date of birth = existing date of birth

And the incoming gender = existing gender then

If length of incoming and existing last names are both greater than or equal to 5 then

If the difference between the lengths of the incoming and existing last names is less than or equal to 3

If the incoming last name = existing last name comparing the strings using the shorter length of the two strings, then add weight for last name

**** How first name should be compared for setting weight.**

If incoming first name = existing first name, then add weight for first name

Else

If the incoming last name = existing last name

And the incoming date of birth = existing date of birth

And the incoming gender = existing gender then

If length of incoming and existing first names are both greater than or equal to 5 then

If the difference between the lengths of the incoming and existing first names is less than or equal to 3

If the incoming first name = existing first name comparing the strings using the shorter length of the two strings, then add weight for first name

If a single potential match record is found with a match weight of >90

If the incoming record already contains an STI

Match error has already been established by the above processing of an incoming STI, continue with the next record

Otherwise,

Update Record on Student Master

Insert Change History

Increment update student count of File Log

Otherwise if multiple potential match records are found with a match weight of > 90

Set error condition to "high potential match"

Increment resolution count on File Log

Insert record into the Ambiguous Match table

Insert record into the Potential Match table linking the incoming record with the each

Student Master record with a weight > 90

Discard all other potential match records with weight <= 90

Otherwise if one or more match weight of > 81 found

Set error condition to "medium potential match"

Increment resolution count on File Log

Insert record into the Ambiguous Match table

Insert record into the Potential Match table linking the incoming record with the Student Master record with a weight > 81

Discard all other potential match records with weight <= 81

Finally assume a new record

Generate a new Unique Student Identifier

Insert Record into Student Master

Insert Change History

Increment new student count of File Log

The matching weights outlined in this document are putting a higher value on fields that have a greater probability of discriminating matching records. Gender has two characteristics, it is a high accuracy key entry field (only has two possible values) and if a mismatch occurs, this is a high degree of a non-matching record. Subsequently DOB is probably less accurate from a key entry standpoint but by its nature has a higher discrimination value.